Инциденты и коммуникация: что, как и зачем говорить, когда все сломалось

Виктор Попов





X5 RETAIL GROUP

/Виктор Попов

Техлид DevOps-команды департамента по развитию и эксплуатации корневых систем в X5 Retail Group

Помогаю поддерживать и развивать k8s-кластеры as-a-service для 20+ продуктовых команд. А так же кучу сопутствующих инструментов

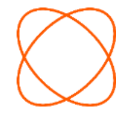
Чиню коммуникации между dev- и ops- командами

Рассказываю пользователям, чем вообще занимаются инженеры

Бегаю по разным людям с безумными идеями и мешаю им работать



tg: @IvanovIvanIvanovich1
mail: vik.popov@x5.ru





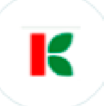
X5 RETAIL GROUP

/О КОМПАНИИ

X5 Retail Group — ведущая продуктовая розничная компания России.

Управляет брендами:














X5 ready food

DIALOG X5



25 лет

На рынке продуктового ритейла в России

17 959

Магазинов*

12,8%

Доля рынка Продуктовой розницы

12,6%

Доля онлайн-рынка Продуктов питания

~ 340 тыс.

Сотрудников*

*Данные на 31 марта 2021 года





/О X5 ТЕХНОЛОГИИ

X5 Технологии — бизнес-единица X5 Retail Group, которая создаёт комплексные
цифровые решения для бизнеса и торговых сетей компании.

~ 2 700

Сотрудников

34 / 113

Активных продуктов / проектов

< 56 млн

Золотых записей клиентов

5.0 Пб

Объемы хранения кластера больших данных

324 / 20

Ит-систем в эксплуатаци/критичных

> 1 200

Физических серверов







/ГРАНИЦЫ ПРИМЕНИМОСТИ

Все это мое ИМХО











/ГРАНИЦЫ ПРИМЕНИМОСТИ



Все это мое ИМХО

И опыт









/ГРАНИЦЫ ПРИМЕНИМОСТИ

Невраждебные пользователи

Технически грамотные







ЧАСТЬ 1

ПРО ДОВЕРИЕ



«TRUST ME»

BT 7274 TITANFALL 2





/SHIT HAPPENS

НЕ ЕСЛИ, А КОГДА

OVHcloud

HighLoad++
Весна 2021



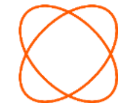
/ПРОЗРАЧНОСТЬ

Пользователи узнают о проблеме



«What we don't understand we can make mean anything. »

Chuck Palahniuk «Diary» (2003)





/ПРОЗРАЧНОСТЬ

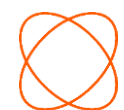
Пользователи узнают о проблеме

Лучше если они узнают от вас



«What we don’t understand we can make mean anything. »

Chuck Palahniuk «Diary» (2003)







/ПРОЗРАЧНОСТЬ:GITLAB

Все метрики доступны на

публичных дашбордах

ССЫЛКА ТУТ





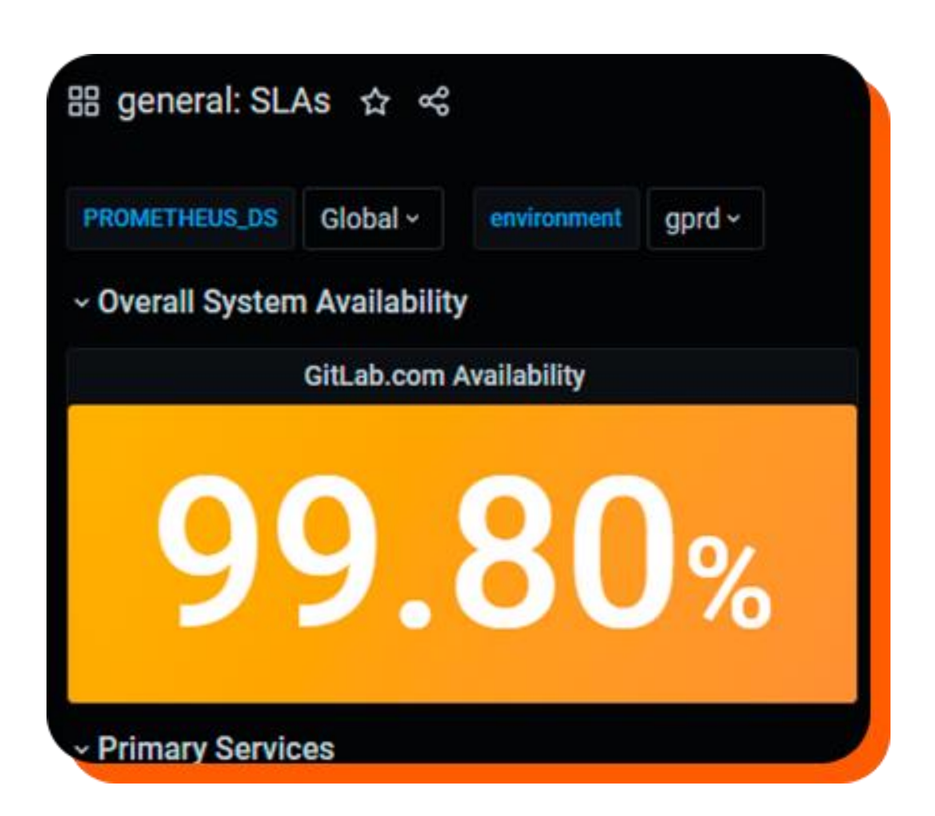




/ПРОЗРАЧНОСТЬ:GITLAB

Даже вот такие:

ССЫЛКА ТУТ



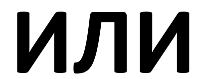


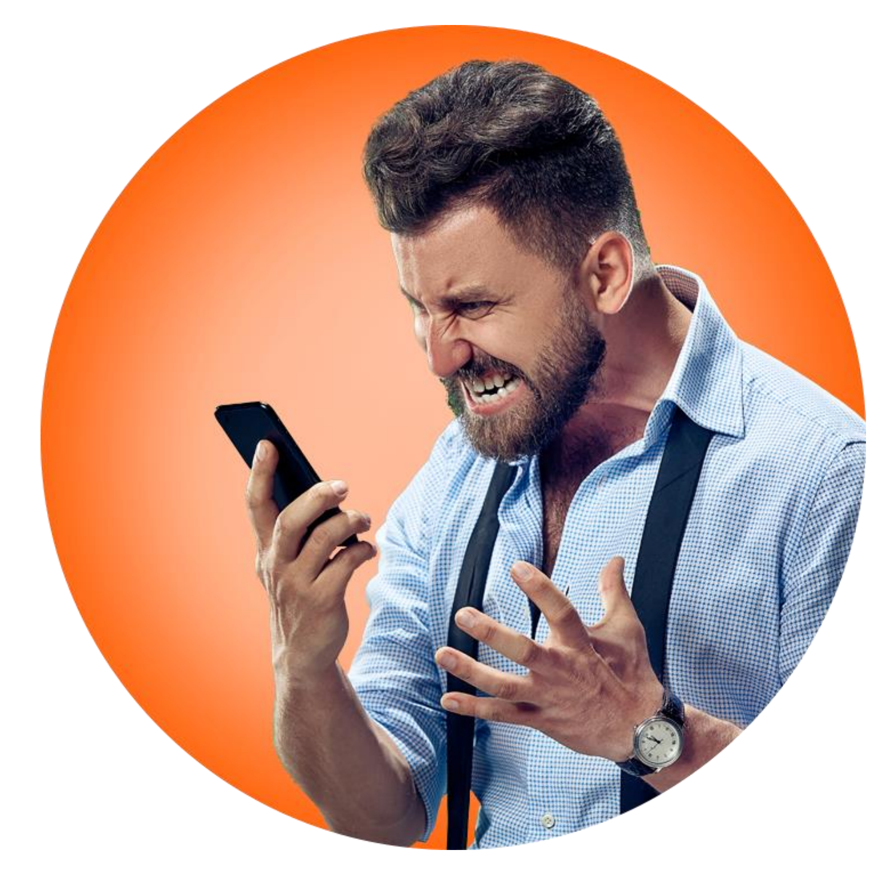




/ВАЖНОСТЬ ДОВЕРИЯ











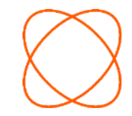




/ЧЕЛОВЕЧЕСКИЙ ФАКТОР

Самая частая причина проблем











/ЧЕЛОВЕЧЕСКИЙ ФАКТОР

Самая частая причина проблем

Не существует*









/ПРОБЛЕМУ МОЖНО

Скрыть и не сделать выводы (но это прям фу таким быть)





/НЕ ВРАТЬ
И НЕ БОЯТЬСЯ

Скрыть и сделать выводы

Рассказать и сделать выводы



«I never lie because I don’t fear anyone. You only lie when you’re afraid »

John Gotti American Criminal



ЧАСТЬ 2

ПРО ИНЦИДЕНТЫ



**«IT'S SO HARD TO FORGET PAIN, BUT IT,S EVEN HARDER TO REMEMBER SWEETNESS.
WE HAVE NO SCAR TO SHOW FOR HAPPINESS.
WE LEARN SO LITTLE FROM PLACE»**

CHUCK PALAHNIUK





/РОЛИ. ПОЖАРНЫЙ

Тот, кто прямо сейчас всех спасает











/РОЛИ. КООРДИНАТОР ПОЖАРА

Держит общую картину

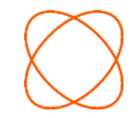
Транслирует решения команде

Распараллеливает задачи

Сообщает о взаимомешающих задачах

Ведет лог инцидента









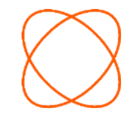


/РОЛИ. КООРДИНАТОР ПОЖАРА

Не обязательно самый опытный или высокий по должности

Не бойтесь взять эту роль







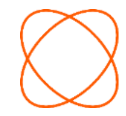




/РОЛИ. ЛИДЕР КОММУНИКАЦИИ

Транслирует происходящее









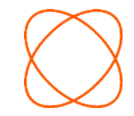


/РОЛИ. ЛИДЕР КОММУНИКАЦИИ

Защищает пожарных

Особенно от больших боссов

HOLD THE DOOR









/РОЛИ. ПЕРЕДАЕМ ЭСТАФЕТУ

Важно ЯВНО передать роли, задачи, результаты

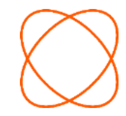
Здесь поможет лог инцидента

Героизм может сделать хуже

Отдыхать нужно

Инженеров нужно кормить!







X5 RETAIL GROUP



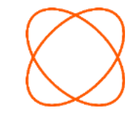
/РОЛИ

Должны быть известны и понятны
Границы ответственности установлены заранее
Степень формальности – на ваш вкус

Вот как это сделано у гитлаба:

During an incident, the CMOC will:

1. Be the voice of GitLab during an incident by updating our end-users, internal parties, and executive-level managers through updates to our status page hosted by Status.io.
2. Update the status page at regular intervals in accordance with the severity of the incident.
3. Notify GitLab stakeholders (e-group, customer success and community team) of current incident and reference where to find further information. Provide additional update when the incident is mitigated.









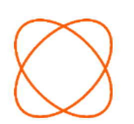
/КОММУНИКАЦИИ ВНУТРИ КОМАНДЫ

Должна быть обговорена заранее

Мессенджеры оставляют артефакты

Если у вас war room – нужен летописец

Доверие внутри команды







/ВО ВРЕМЯ ИНЦИДЕНТА

Не молчим

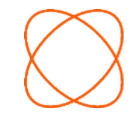
Не флудим

Регулярный статус

Доносим информацию только тем, кому она нужна

Трезво оцениваем сроки











/УРА, МЫ ВСЕ ПОЧИНИЛИ

Сообщаем пользователям:

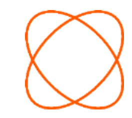
Статус

Ограничения в работе

Что нужно сделать с их стороны

Когда вернется в 100% рабочее состояние









/POSTMORTEM

Тут пригодится лог инцидента и координатор

Хронология

Откуда узнали

Что произошло, что делали, что не помогло, что помогло

Что сделать, чтобы такого больше не было?



Не советую гуглить postmortem
в поиске изображений.

Поэтому тут белочка









/POSTMORTEM.
5 «ПОЧЕМУ»

Problem 1: GitLab.com was down for about 18 hours.

1. **Why was GitLab.com down?** - The database directory of the primary database was removed by accident, instead of removing the database directory of the secondary.
2. **Why was the database directory removed?** - Database replication stopped, requiring the secondary to be reset/rebuilt. This in turn requires that the PostgreSQL data directory is empty. Restoring this required manual work as this was not automated, nor was it documented properly.
3. **Why did replication stop?** - A spike in database load caused the database replication process to stop. This was due to the primary removing WAL segments before the secondary could replicate them.
4. **Why did the database load increase?** - This was caused by two events happening at the same time: an increase in spam, and a process trying to remove a GitLab employee and their associated data.
5. **Why was a GitLab employee scheduled for removal?** - The employee was reported for abuse by a troll. The current system used for responding to abuse reports makes it too easy to overlook the details of those reported. As a result the employee was accidentally scheduled for removal.





/POSTMORTEM

Постмортем без реальных задач – плохой постмортем

Вот тут пример отличного постмортема

от гитлаба:

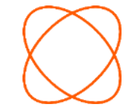
ССЫЛКА ТУТ



Не советую гуглить postmortem

в поиске изображений.

Поэтому тут белочка







/BLAMELESS POSTMORTEM

Здоровая культура в компании

Снижает вероятность инцидента из-за страха признаться в ошибке

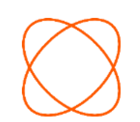
Вот тут ребята из Atlassian классно рассказывают про blameless в постмортемах:

ССЫЛКА ТУТ



Не советую гуглить postmortem в поиске изображений.

Поэтому тут белочка







/СООБЩАЕМ ИТОГИ ИНЦЕДЕНТА

Простым языком

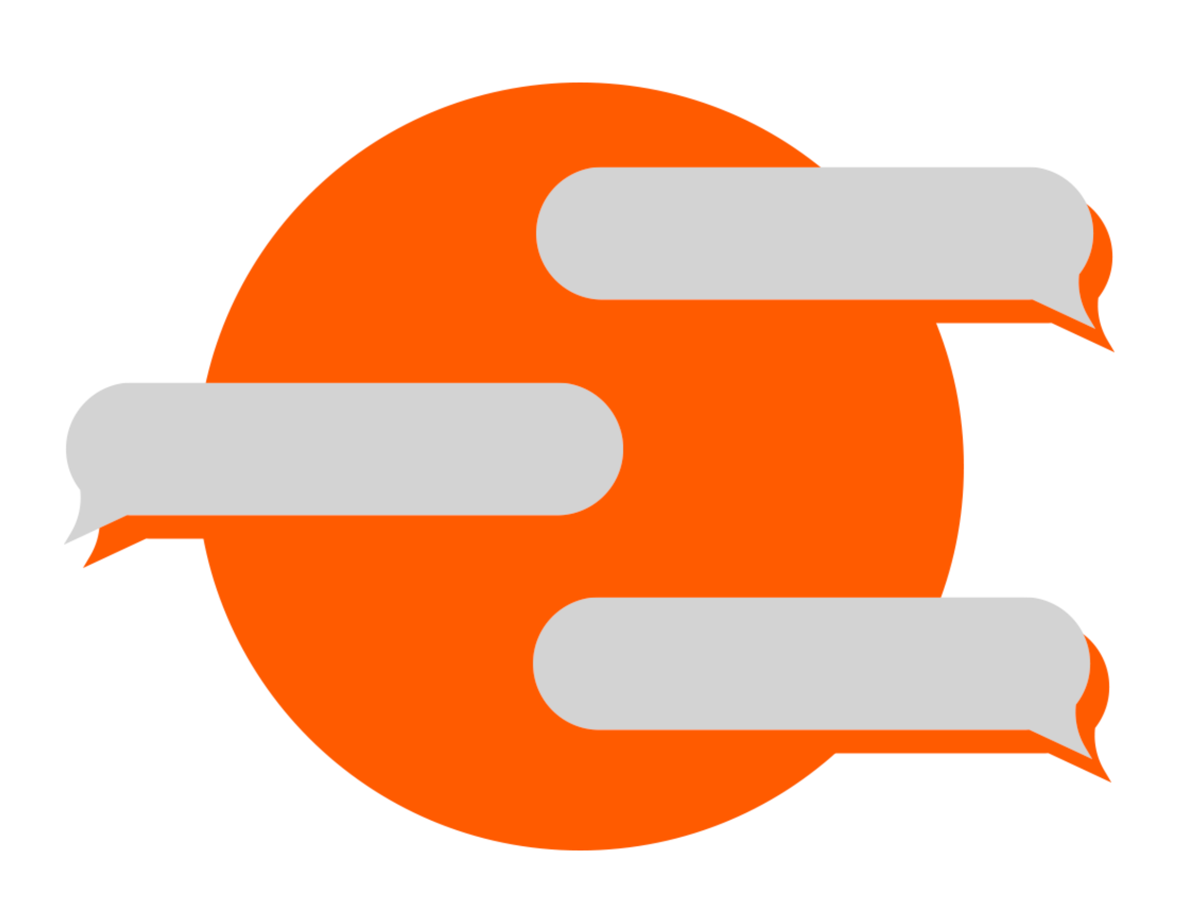
Кратко

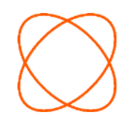
Почему

Как починили

Почему не повторится

Ссылка на подробный анализ в постмортеме









/КАНАЛЫ КОММУНИКАЦИИ

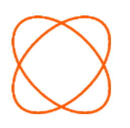


Нет плохих, есть не подходящие

Используйте то, что удобно вашим пользователям, а не вам

Даже ради нескольких пользователей стоит использовать еще один канал связи





ЧАСТЬ 3

ЧТО ЕЩЕ ПИСАТЬ?

«THE KEY TO
A GREAT STORY
IS NOT WHO,
OR WHAT, OR WHEN
BUT WHY.»

ELLIOT CARVER







/ПИШИТЕ НЕ ТОЛЬКО ПРО ИНЦИДЕНТЫ!

Коммуникация ведёт к доверию

Пользователи узнают, чем вы вообще заняты и почему их тикеты не решаются мгновенно

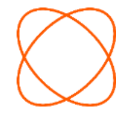
Пользователи научатся чему-то новому

Меньше одинаковых вопросов



Я к вам пишу – чего же боле?
Что я могу еще сказать?
Теперь я знаю в вашей воле
Меня презреньем наказать.

А. С. Пушкин







/ЧТО ПИСАТЬ. АНОНСЫ РАБОТ

С человеческим лицом

Объясняем зачем

Обосновываем срочность



Когда?
22 декабря, 20:00 -- 23:00

Зачем?
Исправить баг в s3, не позволяющий снимать бекапы.
Так же появится функционал ресайза cephfs вольюмов.

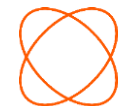
Срочность
От обновления зависит еще несколько важных внутрипродуктовых и инфраструктурных задач, которые очень хотелось бы сделать в этом году. KPI горит, классика :)
Поэтому очень хотелось бы иметь запас по времени, если во время обновления снова возникнут сложности.

Влияние
Потушим ранчер и гитлаб, чтобы быть увереными, что никто не создаст новые volumeclaim.

Так же вероятно снижение производительности стораджа, но в прошлый раз жалоб замечено не было.

601 edited 11:53











/ЧТО ПИСАТЬ. НОВЫЙ ФУНКЦИОНАЛ

«Оправдывает» обновления

Пользователи узнают про новые фишки

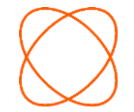
Обновили gitlab с 13.0 до 13.2. Что нового?

13.1
Аналитика конрибуторов теперь показывает количество принятых MR.
Фильтры пайплайнов по статусу и тегу.
Переменные CI/CD уровня инстанса gitlab.
Шаблоны для rules. Подробности тут:
https://docs.gitlab.com/ee/ci/yaml/index.html#workflowrules-templates
Добавили поиск по группе.

Полный список изменений доступен тут:
https://about.gitlab.com/releases/2020/06/22/gitlab-13-1-released/

13.2










/ЧТО ПИСАТЬ. АНОНС ВАЖНЫХ ИЗМЕНЕНИЙ

Чтобы к вам не прибежали, когда сломается

Сюда же про дорожные карты и планы на будущее

Смена имени ветки по-умолчанию

В git версии 2.31.0 ветки по-умолчанию меняют с master на main.

До gitlab это изменение докатится в 14.0, но уже с 15 марта после удачного apt upgrade вы получите ветку main после локального git init.

В gitlab 13.11 (конец апреля) появится возможность задавать имя ветки по-умолчанию для новых репозиториев и мы планируем выставить там master, чтобы сохранить привычное именование.

Думаю, что важно обговорить naming convention внутри каждой команды отдельны.
Мы не настаиваем на каком-либо из вариантов и оставляем выбор полностью за вами.

650 14:37









/ЧТО ПИСАТЬ. FAQ

Если вас о чём-то спросили больше чем 2 раза – пора ответить массово

Если что – потом будете просто пересылать сообщение

Как настроить сборку gradle через кеширующий прокси?

В build.gradle в блок repositories добавляем

```
  repositories {
      maven { url 'https://            .ru/repository/repo.spring.io/' }
      maven { url 'https://            .ru/repository/maven-central/'
}
      maven { url '
https://            .ru/repository/GradlePluginPortal/' }
      maven { url 'https://            .ru/repository/JCenter/' }
  mavenLocal()
```

В settings.gradle добавляем

```
pluginManagement {
  repositories {
      maven { url '
```



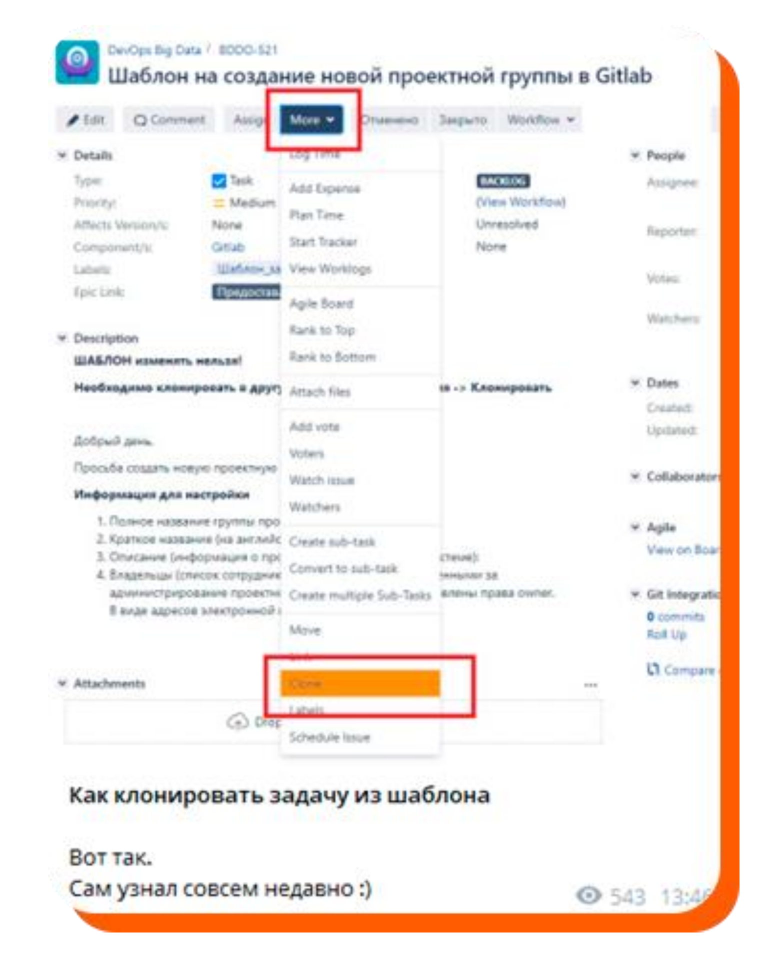


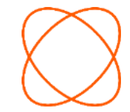




/ЧТО ПИСАТЬ. ПОЛЕЗНЫЕ СОВЕТЫ

Почти как FAQ, только вас еще не спросили, а вы уже всем ответили







ЧАСТЬ 4

КАК ПИСАТЬ?

«КРАТКОСТЬ - СЕСТРА ТАЛАНТА.»

А. П. ЧЕХОВ







/КАК ПИСАТЬ? КРАТКО!

Было

Информирование о работах и сбоях на платформе

Коллеги, мы проводим информирование о событиях на системах и компонентах Платформы больших данных по электронной почте.

Письма рассылаются централизованно на групповые списки рассылки, разделенные по кластерам:



Новости и полезная информация
С новостными сообщениями вы всегда можете ознакомиться в наших телеграмм каналах. Список и ссылки подключения к ним в основном ТГ канале и по ссылке, указанной выше.

1:12

Стало

Информирование о работах и сбоях на платформе

Мы сообщаем о событиях на системах и компонентах Платформы больших данных по электронной почте.

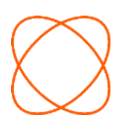
Письма рассылаются группы рассылки, разделенные по кластерам:



Новости и полезная информация
Еще у нас есть информационные телеграм каналы по платформам. Ссылки на подключение тоже есть в вики.

1:12











/КАК ПИСАТЬ? СТРУКТУРИРОВАННО!

Было

Коллеги, мы сделали канал с алертами hive запросов и изменения количеста файлов и места в hdfs.
наш канал: https://t.me/joinchat/placeholder
описание алертов:https://wiki.ru/pages/placeholder
Также помимо самого канала мы подготовили два дашборда которыми вы можете пользоватся
https://kibana.ru/goto/placeholder
https://grafana.ru/d/placeholder
для доступа в дашборд kibana:
login: placeholder
pass: placeholder
тут можно посмотреть будет все данные которые мы получаем позапросу
Мы надемся что данный иструмент будем полезен а так-же ждем от вас обртаной связи.
ваши предложения вы можете оставить в виде коментария вот тут: https://wiki.ru/pages/placeholder
или написать в чат по мониторинга вот тут:
https://t.me/joinchat/placeholder
В скором времени так-же бедут алерты и дашборд по spark в кластере hadoop
edited 1:19

Стало

Публичный мониторинг использования Hive и HDFS

Мы хотим сделать работу с кластером Hive прозрачной, чтобы все пользователи понимали, как их запросы могут влиять на другие команды.

Для этого были созданы
канал с алертами: https://t.me/joinchat
дашборд в kibana с детализацией запросов: https://kibana- ru/goto/
дашборд в grafana с детализацией потребляемого места и количества файлов: https://grafana ru/d/

Если нет доступа в кибану, то можно войти под
login:
pass:

Тут описано, что за алерты и что показывают дашборды:
https:// u/pages/viewpage

Данные инструменты помогут вам оптимизировать запросы и эффективно использовать ресурсы кластера.

Задать вопросы по дашбордам и алертам, а так же что-то предложить можно:
в чате по мониторингу:
https://t.me/joinchat/
В комментарии к статье на вики

В планах аналогичные решения для Spark.
156 14:47









/КАК ПИСАТЬ?
ИНФОРМАЦИОННЫЙ СТИЛЬ

Лаконично	**Было:** В целях повышения доступности сервиса осуществлена миграция на высокодоступную и отказоустойчивую СУБД	**Стало:** Переехали на БД в HA конфигурации
Интересно	**Было:** Произведем работы по повышению стабильности платформы	**Стало:** Переехали на HA базу данных, добавили ресурсов сервису отрисовки котиков, исправили багу XXX.
Честно	**Было:** Сервис предоставляется с ограничениями в течение ближайшего времени	**Стало:** Сломалось создание новых репозиториев. Починим часа за 2







/КАК ПИСАТЬ?

ИНФОРМАЦИОННЫЙ СТИЛЬ

Вот тут подробнее:

ССЫЛКА ТУТ

А вот тут можно проверить текст на читаемость:

ССЫЛКА ТУТ



ЧАСТЬ 5

К ЧЕМУ ВСЕ ЭТО?





/СПАСИБО ЗА ВНИМАНИЕ!

